Technical Report 1658
July 1994

Human Performance Studies for Control of Multiple Remote Systems

Steven A. Murray

Technical Report 1658
July 1994

Human Performance Studies for Control of Multiple Remote Systems

Steven A. Murray

NAVAL COMMAND, CONTROL AND
OCEAN SURVEILLANCE CENTER
RDT&E DIVISION
San Diego, California 92152-5001

K. E. EVANS, CAPT, USN
Commanding Officer

R. T. SHEARER
Executive Director

ADMINISTRATIVE INFORMATION

This work was accomplished by the author in the Adaptive Systems Branch, Code 531, of the Naval Command, Control and Ocean Surveillance Center (NCCOSC), RDT&E Division (NRaD), San Diego, California 92152-5001 while working in the Department of Industrial Engineering, University of Wisconsin, Madison, Wisconsin. Sponsorship was provided by the Office of Naval Research under program element 0602234N.

Released by
LCDR D. E. Demuth, Head
Adaptive Systems Branch

Under authority of
D. W. Murphy, Head
Advanced Systems Division

ACKNOWLEDGMENTS

I would like to express my sincerest gratitude to Dr. Harold Hawkins of the Office of Naval Research, Washington, DC, for his continued support and encouragement; I would also like to thank Dr. Barrett S. Caldwell of the University of Winconsin, Madison, WI, for his commitment. Finally, I would like to acknowledge the generous technical assistance of Dr. Don Bamber, Gary Gilbreath, and Violette Thomas of the Naval Command, Control and Ocean Surveillance Center, RDT&E Division (NRaD), San Diego, CA.

PS

EXECUTIVE SUMMARY

OBJECTIVE

A study was conducted to examine issues relating to single operator control involving multiple remote sensors or platforms using a simulated industrial security environment as the setting. Interface design issues for operator support in such systems can be formidable, as the inherent task complexity creates significant opportunities for operator confusion and overload. Task complexity created significant opportunities.

RESULTS

The operator task was to designate the number and location of intruders in a simulated building, using video information from remote sensor platforms. The experiment manipulated the number of displays which had to be monitored, event rate, image redundancy and sensor platform mobility. Response time increased significantly for increasing numbers of displays, as expected, but also increased independently for event rate image redundancy and for mobile sensors. Results showed that significant performance penalties may be encountered in multiple platform control, and that these penalties accumulate at seemingly low levels of complexity.

RECOMMENDATIONS

In general, this experiment was successful in its goals of identifying and quantifying selected system design variables and their effects on human-machine performance. The study was also successful in identifying the nature and potential significance of variables needing further study. Further work can now be initiated with a clearer understanding of the effects of stimulus redundancy and operator experience on system performance.

CONTENTS

CONTENTS (continued)

INTRODUCTION

The continuing evolution in sensing, computing, and multi-media display technologies has resulted in a range of sophisticated yet practical systems which rely on remote human control. The use of such systems can avoid the risks of exposing people to extreme environments, such as patrolling in foul weather, or performing hazardous waste removal. Furthermore, it can preclude the expense of transporting people to a work site and of providing support and protection for them while they complete their work. For many applications, such as industrial security or manufacturing process control, it is desirable to require a single operator to control several remote platforms simultaneously, multiplying the performance benefits of human presence in the overall system design. As with many new efforts in technology application, however, human performance with complex systems is all too often assumed by extrapolating from performance with simpler ones, and design proceeds in a naive fashion. An attempt to understand the unique control demands for multiple systems and to establish some initial design guidelines is, therefore, a worthwhile design effort.

Designing for control of multiple remote devices combines the user interface problems of single system operations (e.g., telerobotic manipulation or remote sensor monitoring) with control of large-scale processes (e.g., power plant safety or supervision of automated production processes). In the first application, issues of display quality and control feedback are central to operator effectiveness; the task is fundamentally perceptual in nature. In these tasks, provision of good coupling with the physical environment can help to ensure good system performance. In the second application, issues of cognitive representations and world models are critical to effective understanding of the large data loads involved. In this case, clear meaning is a distinct requirement from clear display. Although human performance research exists which addresses each of these applications (Pepper & Kaomea, 1988; Sheridan & Ferrell, 1963; Woods, 1988), research that integrates these issues is relatively meager.

Engineering and deployment considerations can further challenge human performance when systems are operated in unstructured settings, when signal bandwidths are limited, or when sensor platforms are mobile. Telerobotic devices for undersea and space exploration, and semi-autonomous robots for factory and airport security are examples of active engineering programs that place heavy constraints on data transmission rates while imposing strict standards of robust performance. Too often, these demands of varied operational environments, limited data bandwidths, and platform mobility cascade to the user interface, and the human is left with the job of making the system work. The interface, therefore, is critical to system success. It must allow the operator to efficiently receive information from several remote platforms and to act on that information without adding to confusion or overload in the process.

APPLICATION

This study was motivated by questions concerning expected system performance of new security and industrial monitoring systems which employ autonomous, mobile sensor platforms under the control of a single human operator. Capabilities for independent platform navigation and automatic target detection and recognition have improved to the degree that much of the effort required of human security guards may be reliably replaced with sophisticated machinery that can patrol for long periods at sustained high levels of vigilance. With the addition of

intelligent alerts, human operators can remain in one (protected) location while surveillance is maintained over a wide area. Significant events can trigger warnings at the operator station, accompanied by video imagery and other data regarding the cause of the warning. With few clear guidelines to rely upon, however, designers of such systems have been required to define the operator's role themselves and to hypothesize about expected operator performance. Unfortunately, this design approach is vulnerable to excessively optimistic assumptions regarding the environment and sensor performance, etc. What, for example, might happen if several platforms sense events simultaneously and transmit their alerts together? What if events were correlated, as might be the case if an intruder were detected by different sensors whose fields of coverage just happened to overlap on the same event? A security surveillance experiment was conducted, therefore, to generate human performance data about these and other issues.

The experiment performed here was based on an existing real world job — industrial security. This task, however, is more general and can demonstrate the ability of a human operator to rapidly assess information from a number of distributed sources and to integrate it into a single "picture" of the environment. This is a typical function of many existing systems (e.g., bank security stations, utility plant control consoles, etc.), and therefore the results should have a wider degree of application.

In general, control of current and emerging systems with multiple platform require the operator to perform four functions:

1. The user must monitor and/or coordinate the placement of sensors throughout some space or territory. If these sensors are mobile (and especially if they move autonomously), the operator's job of sensor interpretation is greatly complicated, when compared to fixed-platform systems.

2. The user must determine the orientation and perspective of each active sensor and develop a single spatial model of the depicted information. Responses for the security jobs of interest here are usually made by designating alert points on a map or building diagram. Operators must therefore interpret sensor imagery to the extent that they can correlate displayed scenes to some two-dimensional representation of the environment. Performing this job with views from fixed cameras is at the simplest end of this task spectrum, while correlating images from mobile sensors with 360° panning capabilities is at the most complex. In the latter case, information about the location of a sensor is frequently insufficient to determine the location of displayed objects; where the sensor is looking is critical, as well.

3. Objects (such as human intruders) must be detected and identified in each of a number of displays. These objects may appear in an almost infinite variety of locations, positions, or orientations and must be distinguished from other background features such as furniture, shelves, or equipment. Even with the use of automated detection and alerting, the operator must manually examine the display, if only to confirm a false alarm.

4. The operator must integrate the information from multiple displays to construct a single world view of the environment. Both fixed and mobile platforms may survey overlapping fields of coverage, however, from a variety of perspectives when detection alarms are tripped. Thus, the operator must decide whether multiple objects are distinct entities or whether the same event has merely triggered more than one sensor.

Situation awareness, therefore, requires knowledge of the data from each sensor and the distribution of those sensors in the environment. Establishing and supporting situation awareness, then, is the key objective of interface design.

THEORETICAL FRAMEWORK

A review of perception, cognition, and engineering literature is summarized to provide a basis for describing human performance characteristics in complex systems and to help identify critical system variables which influence such performance.

Three operator tasks are emphasized for the present experiment: the number of displays which must be monitored, the rate of information presentation, and the complexity of the information. The primary task of the operator is to determine whether an event has or has not using multiple visual displays. In practice, the number of displays that can be designed for a system is essentially arbitrary, which means that some systems can have quite large search sets. In addition, the rate of stimulus events in most systems is driven by the environment and not the design. The operator may be faced with low stimulus rates (e.g., a night watchman at an industrial plant) or high ones (e.g., an industrial inspection task). Finally, the system application may or may not constrain the range of signals which must be responded to. Thus, a night watchman might be interested in any unauthorized activity in his area while a military surveillance operator might need to distinguish between a wide variety of signals to identify only those of some specified class.

PERCEPTION PRINCIPLES

Research concerning target detection and localization speed shows that the major influence on search time is the size of the search set (Scanlan, 1977). Such literature predicts that operator performance will diminish with larger numbers of displays, regardless of the content in those displays, because the operator needs more time to scan more images. A maximum of three displays was used in this experiment to confine the time commitment required of subjects. It is important to note, however, that some systems are being designed with far higher numbers of displays and that performance with such systems could be much lower than measured by this study.

COGNITION PRINCIPLES

Signal Detection Theory has been used primarily as a means to measure human sensory capabilities (Wickens, 1984, chapter 2). New research is expanding the application of this approach beyond psychophysical testing, however (O'Brien & Feldman, 1992), by expanding the paradigm definition of noise and increasing the range of applicable tasks beyond those of simple detection. This extension provides useful performance predictions for the control tasks of interest here. Specifically, signal detection theory predicts that multiple sources of information will increase the signal-noise ratio and will, therefore, improve signal detection. Multiple information sources may include multiple perspective views of the same object (e.g., through overlapping fields of view) or through presentation of information from several sensor types (e.g., infra-red, motion detection, etc.). This performance change may take a number of forms, such as reduced response time or increased accuracy of target detection.

The primary form of complexity for this study is information redundancy. Because sophisticated sensors are normal components of many modern security systems, overlapping views of the same target are likely; therefore, multiple common cues will exist across some displays in some circumstances. If two sensor cameras, for example, both viewed a floor lamp next to a book case, then the operator would see two lamps and two bookcases in his two-display suite. Their appearances would be different in each display due to the different camera vantage points. If an operator were familiar with the room and recognized these features as common, then a single percept of the scene would be readily formed. If, however, an operator were not familiar with the room, then additional time and effort might be required to establish the common source of the images. In more complex examples (e.g., warehouses, where many objects have similar forms, such as shelves and doorways), operators might have even greater difficulty in this task. In this case, the multitude of cues might add to the burden of interpretation (i.e., adding to the noise), rather than enhancing it (i.e., by adding to signal strength).

To confine the variables of this experiment, only trained subjects and structured environments were employed, although the issues associated with untrained operators and unstructured settings certainly deserve study.

Information Theory also offers useful insights which are relevant to problems of multiple system control (Ashby, 1956). Operators of remote systems must receive inputs about the environment and about the results of their own manipulations of it to perform their tasks. Both types of inputs are information. A bit of information is defined as logN, where N is the number of alternatives (Shannon & Weaver, 1949). Thus, a display in which the operator had to determine only the presence or absence of a target would contain log(2) = 1 *bit of information*. If the operator had to determine the type of target from among several possibilities (i.e., a classification task), the information contained in the display would be higher.

For this experiment, the task is limited to determining the presence or absence of a human figure in the image of each of a number of displays (i.e., a two alternative task, multiplied by the number of displays). The information theory model of most relevance in this application is the Hick-Hyman law (Hyman, 1953). Briefly, this law states that reaction time (RT) increases linearly with increases in stimulus information, i.e., $RT = \log_2 N$. It is important to note that this law is stated for RT, which is normally defined as the interval from stimulus onset to initiation of a response. Additional time is required to actually execute the response which, when added to RT, constitutes the response time for a trial. Because response time was all that was measured for the present study, the noise introduced from this more global measure necessarily adds noise to the dependent measure. A monotonic increase in response time, however, should nevertheless provide verification of the increasing information processing requirements of the task conditions.

Vickers (1970) adds an important condition, stating that RT goes up as stimuli become less discriminable from each other. Although discriminability would certainly be a useful manipulation, the desire to provide unambiguous data at this stage of investigation constrained all stimuli used in the study to a condition of clear discriminability.

The speed-accuracy tradeoff (Wickens, 1984, chapter 9) is the phenomenon whereby errors tend to increase as people attempt to respond more quickly to task demands. This, too, is explainable in terms of information theory (Hick, 1952), but also helps to explain performance as operators select strategies which will maximize efficiency (Fitts, 1966). Accuracy was

emphasized for this study, and instructions to subjects intentionally biased the participants toward this emphasis. For this reason, the primary performance measure was response time.

Workload is a concept of direct concern to the kind of complex system control described earlier. Workload has been defined as the difference between the resources of the information processing system that are required for task performance and the resources available at any given time (Gopher & Donchin, 1986, chapter 41). Processing demands for multiple platform systems can challenge the capabilities of the operator, imposing a mental workload. The primary thread which connects the various lines of research in this area is the argument for single versus multiple cognitive resources to explain the observed limitations on human ability to process information. The majority of these studies tend to confirm multiple processes (Wickens et al., 1982, 1983; Wickens, 1987) or pools of resources. This additional perspective is useful to the present study because it addresses both the size of the operator display set, as well as the number of displayed targets and their supporting visual cues. Simultaneous manipulation of these variables might be explained in terms of whether distinct or common perceptual resources are required to perform this task.

ENGINEERING PRINCIPLES

To establish and maintain human understanding of the environment, a display interface must map the physical world to the user's mental representation of that world (Wickens, 1984, chapter 5); it must, in other words, maintain a congruence between reality and representation. If the mapping is natural, less cognitive processing is required to update mental representations. If the interface presents information which must be transformed, however, then additional processing is required and the resulting mapping process should be slower or less accurate. Wickens (1984) applies this principle by advocating the most natural interface mapping between reality and representation, i.e., a mapping which preserves both the static aspects of object orientation as well was the dynamic properties of object motion. The multiple platform monitoring task involves frequent changes in orientation and, therefore, lends itself to tests of the display-response compatibility principle. One such test is included in the experiment reported here.

Maps are two-dimensional spatial representations of a three-dimensional world. Thorndyke and Stasz (1979) found individuals with higher spatial ability did better on map-learning tasks than those with lower spatial ability (but equivalent intelligence). Furthermore, the qualitative nature of map knowledge, i.e., forming a cognitive map, undergoes changes due to learning and proceeds through fixed processing stages (Thorndyke, 1980). These results are important because the task of interest here necessarily involves transitions from three-dimensional sensor displays to a two-dimensional depiction of the environment to integrate data from multiple platforms into a single "big picture."

Current operator displays are typically fixed in place (e.g., rows of CRTs), and have historically been adequate for fixed sets of sensors. The effects of such fixed displays with mobile platforms, however, is unknown. Systems of this type have not yet been subjected to field test. What, for instance, is the performance impact of interpreting sensor information when the order of the displays on a console is different from the order of the platforms in the plant? Figure 1 illustrates an example of such a situation. The top half of the figure shows the distribution of three mobile platforms in a building, while their accompanying sensor images are shown in the bottom portion of the figure, displayed on a fixed bank of CRTs. The lateral (i.e., left-right) distribution of

these images on the operator display does not match the lateral order of the sensor platforms themselves, and some mental translation is required to make sense of the world view represented by this information. Furthermore, information about the vertical distribution of the sensors is lacking completely; sensor 2 is "highest" in the diagram, while sensor 1 is "lowest." This information cannot be depicted with the display scheme shown in the figure.

Applying graphical computer interface tools, such as windowing environments, might aid the operator's orientation task, especially regarding the physical correspondence between the platform locations and their displayed imagery. A single candidate application is examined in this study, based on the display-task correspondence principle discussed by Wickens (1984). Specifically, this alternative scheme matches the distribution of display images to the distribution of platforms in the building, both laterally and vertically, as shown in figure 2. This approach implicitly treats the display surface as a direct analog of the building layout. The integration of information regarding platform distribution and corresponding imagery should reduce the need for the operator to sequence back and forth between the two types of displays, as some current system designs require. This should consequently ease the short-term memory burden and possibly help to resolve scene ambiguity. Because the display presents images that conform to physical sensor locations, this approach is referred to in the experiment as a conformal display.

Note that for fixed-position sensors, the conformal display would never change and, thus, its relative utility would most likely not be as strong as for a mobile sensor system.

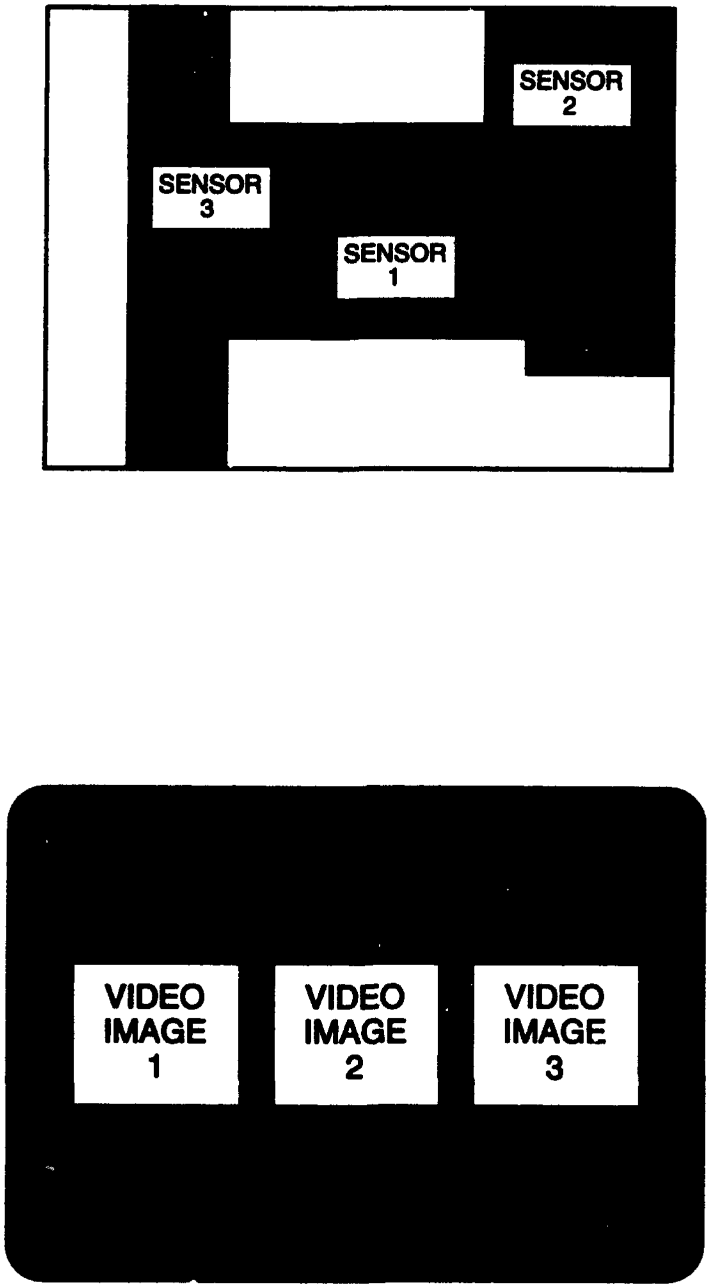

Figure 1. Example of mismatch in lateral position of sensors and operator displays.

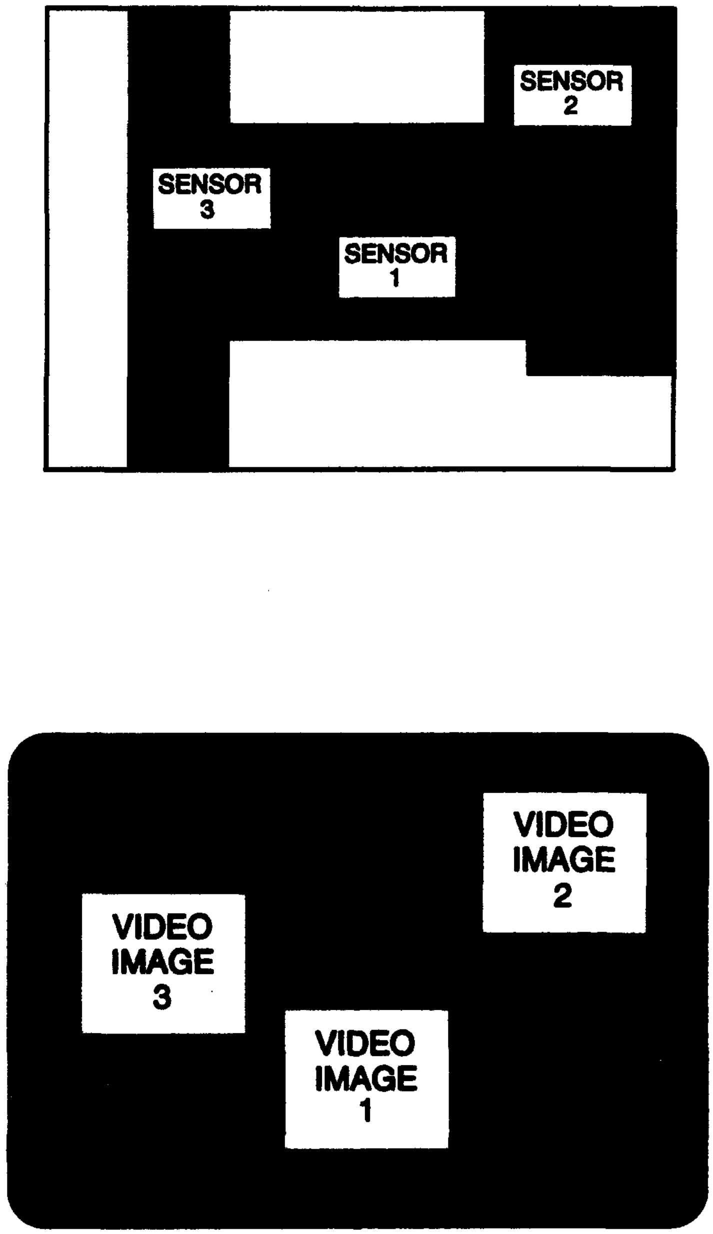

Figure 2. Conformal display concept.

SUMMARY AND HYPOTHESES

The purpose of this study was to determine whether features of multiple platform control could be identified and modeled using existing human factors research principles. Issues such as the number of displays which must be monitored, frequency of events in the environment, sophistication of the sensor platforms themselves (i.e., whether or not they can independently move through the environment), and complexity of the displayed information are formalized here as variables for experimental study, and hypotheses are offered regarding expected outcomes.

Research into visual performance (Scanlan, 1977) has found that a large number of stimuli are more difficult to search than a small number, resulting in increased task performance time, increased error rates, or both. The number of sensors (or the number of displays which must be monitored) therefore has a direct effect on human visual search times and the consequent performance which can be expected of the total system.

Studies of cognitive workload (summarized in Wickens, 1984) have also shown that the greater the stimulus event rate, the greater the perceptual and cognitive workload on the operator. Some events may be infrequent and independent, such as an intrusion into a factory. Other events may be highly correlated and may increase rapidly in number, such as the presence of foot soldiers in advance of an attacking force. The impact of this issue, therefore, may depend on the application of the particular system. For this study, only simultaneous events are considered, i.e., all signals arrive to the operator at the same time. The manipulation of event rate is achieved by varying the number of targets depicted in the images (*described in the* METHOD *section*). Additional examination of this task in terms of the number and rate of sequential alerts, however, certainly deserves research attention.

More complex events should be more difficult to interpret than simpler ones. Many factors can enter into the complexity equation, including the correlation or redundancy among events, the number of different types of events which can occur (i.e., the size of the stimulus set), and the time history of events occurring in sequence. For the vast majority of surveillance applications, "event" and "image" will be synonymous. That is, the presence of an object of interest in a sensor display will constitute an event, and detection or classification of that object will constitute the task. For some applications, however, an event will be defined by much more than physical presence. A particular type of vehicle, for example, or the nature of the (human) activity shown in a display might signal many different consequences than would be indicated by presence alone. Although a rich source of research topics lies in this issue, the study reported here will treat only a single aspect, that of object redundancy.

Mobile, patrolling sensor platforms add the requirement to monitor continually evolving patterns of platform distributions in order to properly interpret sensor imagery. The added flexibility and power of a system of mobile robots can be offset by the increased perceptual and cognitive demands on the operator, especially if the task is designed so that monitoring is infrequent (e.g., if the operator only needs to scan the system displays in response to intrusion alerts).

An earlier discussion cited potential confusions that can result from coupling large numbers of remote platforms, especially mobile platforms, to physically fixed sets of displays. Various alternative display methods exist, which can mediate such confusions, but their relative merits

can usually only be determined after special analysis and testing. The issue of display-task congruence, however, is a legitimate topic for this study and an initial test of this congruence is presented in the form of a conformal display format, as an attempt to realize a performance improvement from this concept of interface design.

HYPOTHESES

As stated earlier, the effective security monitor should not dispatch a response to sensor alerts without some minimum level of confidence that his/her assessment of the number and location of events in the environment is accurate. Subject instructions for this experiment will reflect this emphasis, which means that error rates will most likely not change very much in response to manipulations (i.e., subjects will not act before reaching a threshold level of decision confidence), but response times probably will change significantly. Response time, therefore, is predicted to be the more sensitive dependent measure.

1. Performance time should increase linearly as a function of the number of display, due to a minimum scanning time required for each display. For practical purposes, increased response times or error rates in this experiment are predictable; what is not predictable, however, is the slope of this increase.

2. Response times should also increase linearly as the number of events (i.e., the number of figures shown in the displayed images) increases, according to predictions of the Hick-Hyman law. These response times should be greater than those obtained by simply increasing the number of displays, as this task requires additional visual and cognitive processing to determine whether or not an intruder is present in each image.

3. Redundant figures (i.e., conditions where the same figure appears on more than one display, although from different perspectives) should result in lower response times than nonredundant figures. According to signal detection theory, redundant figures have more confirmatory information (i.e., they strengthen signal over noise). If this interpretation of the theory holds, operators will be able to take advantage of these cues to help make their judgments. Thus, someone monitoring a familiar area (e.g., a storeroom) could find overlapping views to be of assistance in gaining situation awareness. Someone monitoring an unstructured scene, or one they were unfamiliar with, would find this additional "clutter" confusing, resulting in poorer performance. Because only trained operators are used as subjects in this experiment, they should be familiar with the information that would likely be contained in the display images and therefore should find a performance benefit to redundant information.

Because the possible locations of the sensor platforms are essentially infinite, task performance should be significantly more difficult in the mobile condition versus the fixed condition, resulting in longer response times. Constant mobility and redistribution of these platforms precludes the opportunity to practice and learn sensor patterns or perspectives; orientation must be reestablished for every event.

4. The impact of mobile sensors should manifest itself in the form of diminished operator task performance (i.e., higher response times). The task selected for this experiment requires the operator to identify and integrate display viewpoints that change over time, and should therefore place an additional cognitive burden on the operator when compared to control of a fixed-sensor system.

5. The conformal display condition, i.e., the display which matches displayed images to their horizontal position in the environment, should support better performance in the form of shorter response times. In accordance with Wickens (1984) (and others), the principle of display-response matching implies that any display design measures that increase the congruence between the real world and the operator's mental representation of the world should show improved performance.

METHOD

An indoor setting was selected for this experiment. That is, the surveillance area for the sensor systems was confined to the inside of a single building, as they would be for a factory ... ty application. This constraint made for a more controlled experiment.

EQUIPMENT

A Silicon Graphics Indigo (SGI) computer, with a 43 cm (diagonal), high resolution (1280 x 1024 pixel) monitor, was used as the display system. The internal clock of the SGI was utilized for response time measurements, as described in the STIMULUS PREPARATION section. Diagrams of the environment patrolled by the sensor platforms were constructed on paper, using a separate sheet for each trial. These diagrams were approximately 9 x 12 cm in size and showed only the basic walls and hallways of the building. Figure 3 depicts the diagram information used by subjects. The top picture shows the major features of the building model. Furnishings and other interior features were not depicted, but were known to the subjects who participated in the study. Sensor platform locations for the experiment were limited to the shaded areas of the building (i.e., the workbay and hallways) to limit the range of cues available in the display images. The bottom picture of figure 3 shows a sample diagram actually used in the experiment, with typical sensor locations and labels.

STIMULUS PREPARATION

A computer animation model had already been created of a building interior used for robot development at the Naval Command, Control and Ocean Surveillance Center (RDT&E Division), in San Diego, CA. This model had been created with Silicon Graphics software and contained an accurate depiction of essentially all features (e.g., doors, furniture, machinery, etc.) found in the actual building. Furthermore, the software supported easy modification, allowing the insertion of simulated "intruders" into the images as test targets.

The experimental stimuli consisted of a set of monchromatic images representing static video scenes, which might be generated by sensor cameras. Images were generated by taking random scenes from the computer model using a screen capture tool. Each of the test stimuli was approximately 9.0 x 7.5 cm in size, and was displayed on a *Silicon Graphics monitor.* A human figure was constructed using these same modeling tools and was placed in a subset of the images as a test target. Each trial was created by generating and placing the appropriate number of intruders in the model (i.e., by "cloning" this basic figure, as necessary), and then establishing the desired number and positions for the sensor platforms. The image from each of these "notional" platforms was then captured and combined into a computer file, which would generate the appropriate combination of images using a shell script.

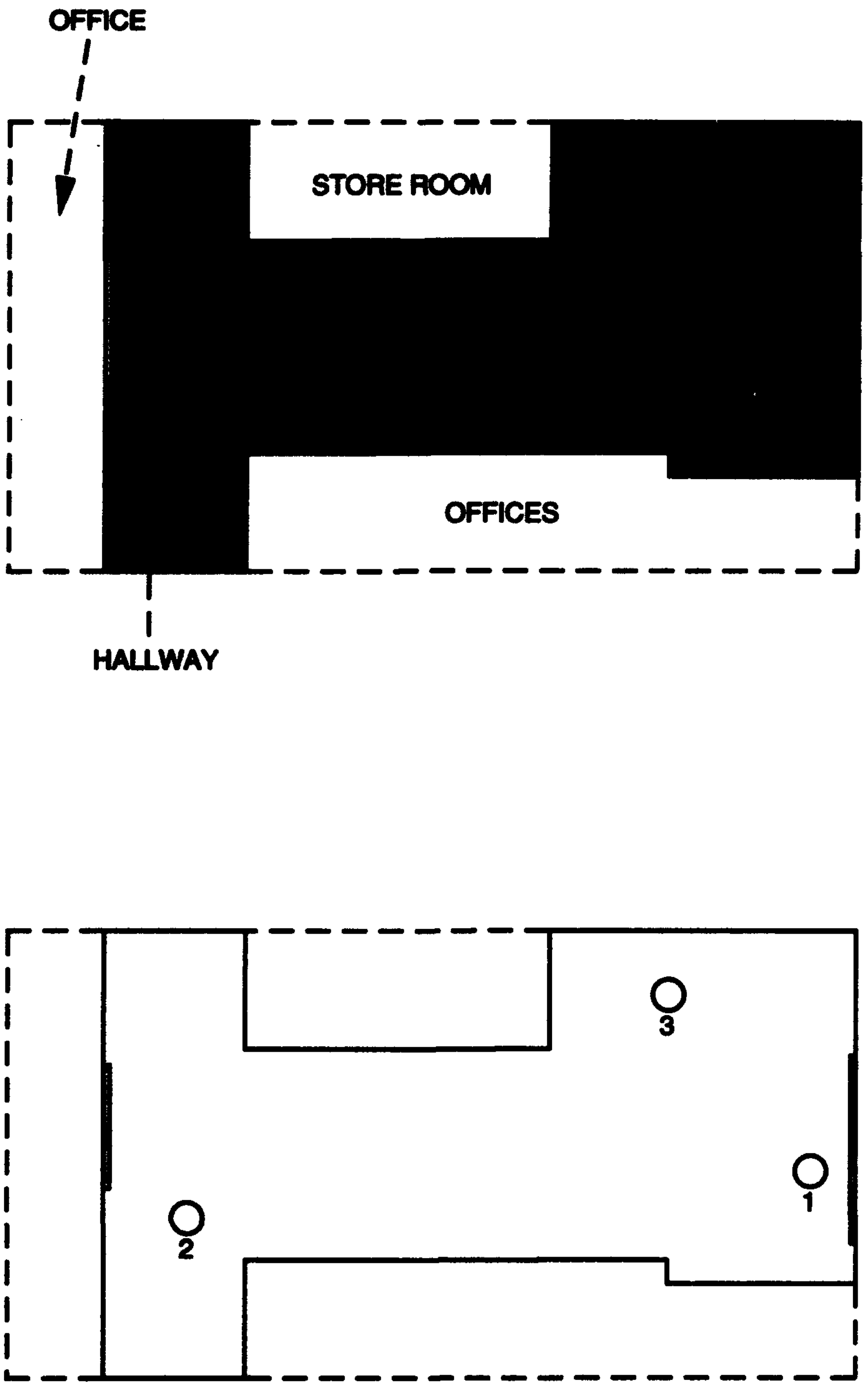

Figure 3. Building layout diagram used by subjects.

Building diagrams are currently available to operators of most existing security systems in the form of charts or maps. Furthermore, many systems include special displays for depicting sensor locations in those buildings, or include them on the charts themselves. To simulate this condition, a building diagram was constructed for each trial, which showed the locations of the sensor platforms. These diagrams were used as a means for subjects to correlate the video images into an overall "big picture" of the environment (see figure 3), and for recording their responses.

Examples of simulated sensor images are shown in figures 4a and 4b. The accompanying diagrams depict the location of the sensor platform that would transmit the associated image. Note that subjects would only have a diagram equivalent to the bottom picture of figure 3 for the actual experiment.

SUBJECTS

Based on pilot study results, it was decided that the first experiments should be conducted only with people who were already familiar with the sensor environment, i.e., people who had spent enough time in the building to readily recognize its visual cues. Thus, six volunteer staff members of the robot lab, four males and two females, were used as subjects. This level of familiarity corresponds to many "real world" situations where security guards usually become acquainted with features of the territory they are expected to monitor. Subject ages ranged from 26 to 32, and all held degrees in engineering or computer science.

DESIGN

Subject availability was an unfortunate constraint for this experiment, so a low-n, within-subjects design was chosen by practical necessity. Because available literature did not predict the relative sensitivity of the variables of interest, only a few modest levels of each manipulation were used. That is, without precise data from earlier work, there was a concern that subjects could be quickly driven to failure by the experimenter unwittingly making conditions too difficult. For these reasons, the following manipulations and levels were included in the study:

Number of Displays. Operators were provided with one, two, or three displays, corresponding to one, two, or three sensor platforms.

Number of Displayed Figures. This is primarily a workload manipulation, tested by varying the number of intruders present during any alert condition. One important qualification was imposed for this experiment, however, as follows. There was no desire to have subjects conduct detailed searches of displays to find and count multiple targets in a single scene; these behaviors were not the focus of this study. For the same reason, it was undesirable for subjects to spend excessive time examining images for intruders that weren't there. Subjects were therefore told that each display would contain, at most, only a single figure. In addition, subjects were told that there would be no false alarms; every trial would contain at least one legitimate intruder in the image set. A single-display system, for example, would contain only one figure; a two-display system could contain either one or two figures; and, a three-display configuration could contain one, two, or three figures. The objective of this manipulation was to examine the effects of integrating possibly ambiguous information, uncluttered by additional complications such as false alarms (although the signal detection manipulations contained in such a scenario might be of interest for future research).

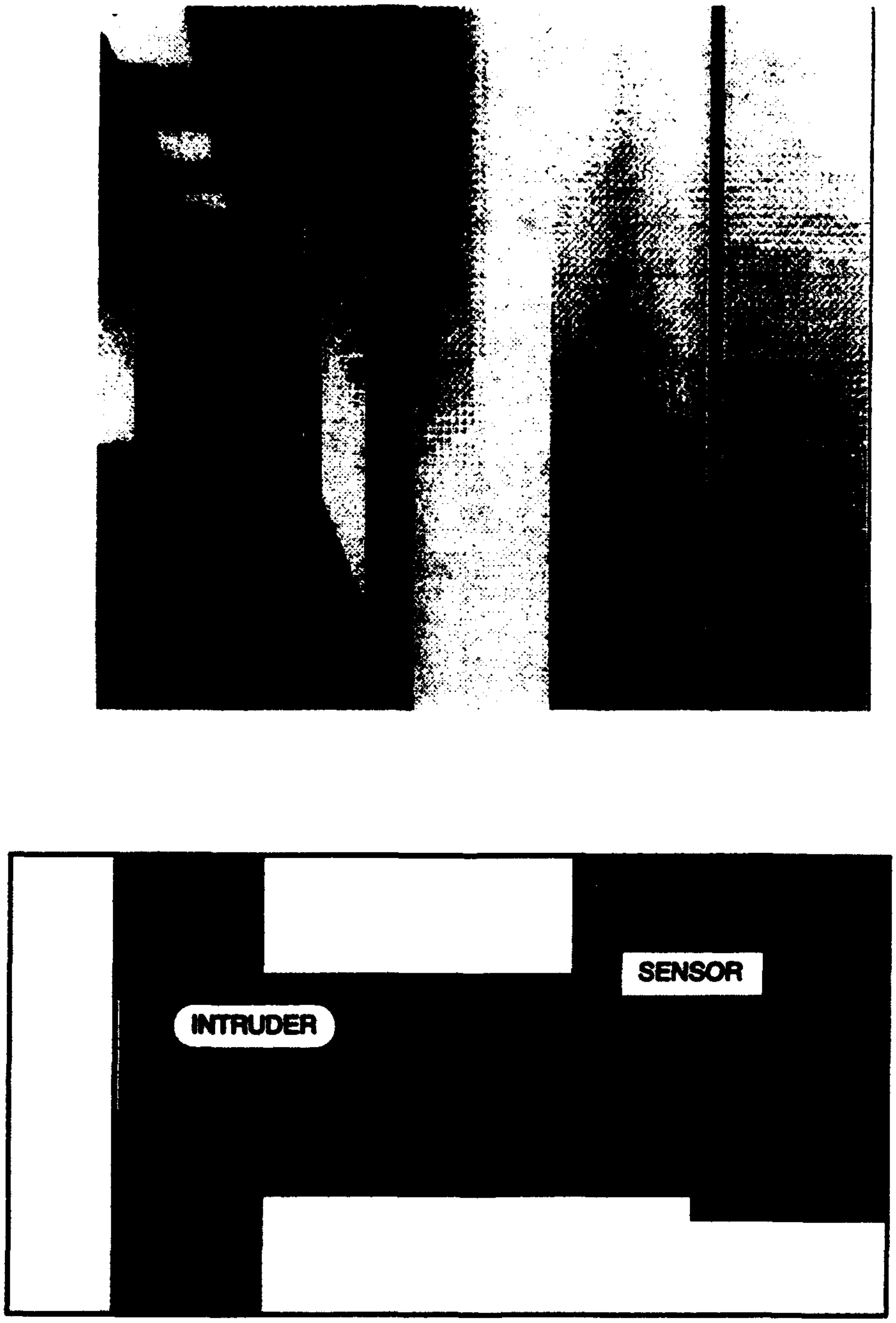

Figure 4a. Examples of sensor image and corresponding location in building.

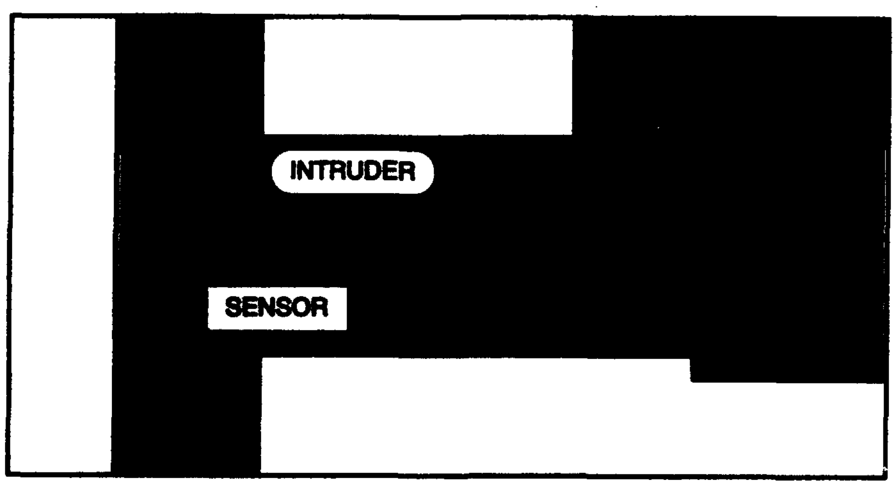

Figure 4b. Examples of sensor image and corresponding location in building.

Number of Actual Figures. This manipulation addressed the kind of image redundancy that might occur when sensors have overlapping fields of view. As discussed earlier, several remote systems could be triggered by the same intruder if they were in physical proximity and their fields-of-view overlapped. An example of this situation is shown in figure 5, where a valid image appears in each display, yet the true situations depicted differ significantly in each case. A trial with three displays, for example, each showing a human figure, might actually represent a situation of three intruders (if all sensors are looking in different locations), two intruders (if two of the three sensors are pointed in the same place), or only a single intruder (if all sensors overlapped). In this case, integration of displayed information is more complex and an operator might waste time or resources in an inappropriate response. It was part of the operator task of integrating these signals to determine the number, as well as the location, of all displayed intruders.

Sensor Mobility. Fixed sensors were defined as those whose physical locations within the building were constant across trials. Fixed sensors could still pan, however, and could therefore show different scenes from the same location in the building. By contrast, mobile platforms could change both their site and viewing direction between trials.

Display Configuration. The distinction between a fixed display console and a conformal console has been discussed in the INTRODUCTION. All experimental manipulations for this experiment were tested with both a fixed (i.e., conventional) display method and a conformal display.

In summary, the experimental design consisted of three levels for number of displays, three levels for number of displayed figures, and three levels for number of actual figures (i.e., scene redundancy, or complexity). In addition, two conditions for sensor mobility and two conditions for display configuration were included. This generated a 3x3x3x2x2 design of 108 cells. The restriction of only permitting one figure per display (described in the Number of Actual Figures section) resulted in only 40 cells that were actually used. Ten trials were conducted for each cell, for a total of 400 trials per subject. The experimental design is summarized in table 1.

PROCEDURE

Each subject was briefed on the purpose of the experiment and told to imagine themselves as a security monitor or surveillance operator. Operator performance in real-world tasks of this type must reflect a balance between rapid and accurate situation assessment. Obviously, rapid response is desirable, but if the situation assessment is wrong, then time may be wasted while erroneous alarms are corrected or personnel are recalled and redeployed. To set a realistic context for the study, therefore, the instructions which were read to subjects emphasized the accuracy criterion.

After an introductory set of instructions had been read, subjects were given a tour of the laboratory where all major visual cues (e.g., windows, doors, tool benches, etc.) were pointed out to them. Orientation to the task was conducted with an initial set of twelve training trials that showed the subject one example of each major experimental manipulation (e.g., fixed vs mobile platforms, fixed vs conformal displays, different numbers of displays (only up to two), and different levels of redundancy in the displayed figures). Subjects were given paper diagrams of the laboratory for each trial and allowed to make their determinations of where the observed figures actually were. Errors, if any, were pointed out for each trial before proceeding to the next one.

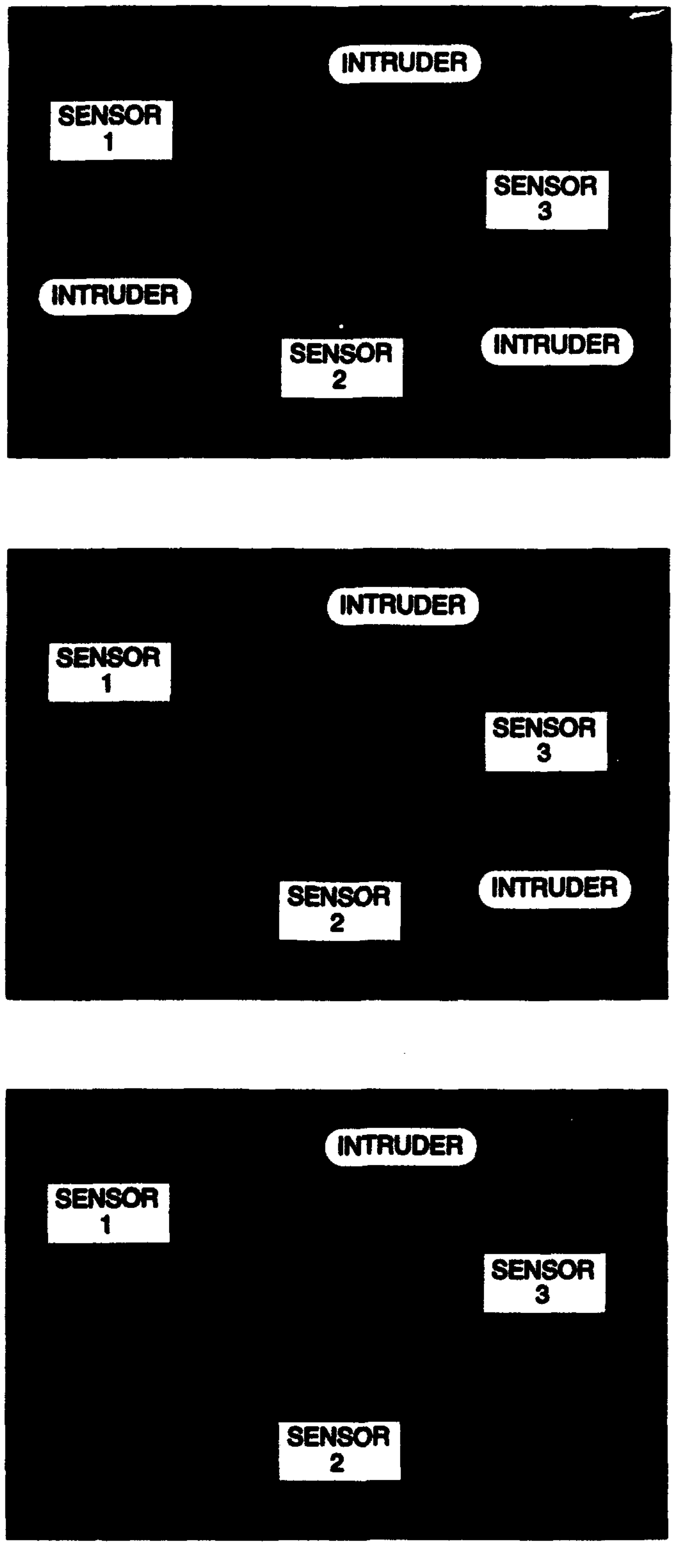

Figure 5. Effect of sensor overlap on displayed objects.

Table 1. Experiment design.

		Fixed Platforms			Mobile Platforms		
		Number of Figures			Number of Figures		
	Number of Displays	1	2	3	1	2	3
Conformal Displays	3	[1]	[1] [2]	[1] [2] [3]	[1]	[1] [2]	[1] [2] [3]
	2	[1]	[1] [2]		[1]	[1] [2]	
	1	[1]			[1]		
Conventional Displays	3	[1]	[1] [2]	[1] [2] [3]	[1]	[1] [2]	[1] [2] [3]
	2	[1]	[1] [2]		[1]	[1] [2]	
	1	[1]			[1]		

Numbers in brackets indicate number of <u>actual</u> figures present in display.

A trial began when the images were shown on the monitor, and ended when the subject indicated (verbally) that they had finished their response. At that time, the experimenter would make a keyboard entry on the Silicon Graphics system to stop the trial, and the system would record the elapsed time since the images had first been presented on the monitor. A response consisted of marking the diagram with the locations — and therefore the numbers — of all intruders that they believed to be in the building (e.g., by placing an 'X' or other mark at the appropriate location on the diagram). During the training sequence, any errors were pointed out to the subject before proceeding to the next trial. No feedback was provided, however, during the experimental sessions.

During the experiment, subjects were always told the system conditions under which they were functioning, i.e., operators knew whether they would be viewing images from fixed or mobile platforms and whether they would be faced with one, two, or three sensors. Within each condition of display number, however, trials were completely randomized for number of displayed figures and the number of actual figures.

DATA ANALYSIS

Each subject received ten trials in each condition, and response time and response accuracy were recorded. Times were compiled using the Silicon Graphics software, while accuracy data were determined by hand scoring each diagram of the experiment. Accuracy data included the number of counting errors (i.e., misjudging how many figures were actually in the building), location errors (i.e., designation of intruders in locations that were off laterally or vertically by *more than 25% of the maximum dimension* of the diagram), and total errors (i.e., a sum of these two error categories) per cell. Data were averaged within cells for each subject and analyzed using a repeated measures ANOVA. Significance was accepted at the $p < .05$ level or better.

RESULTS

DATA SUMMARY

Analyses of response time data revealed statistically significant effects for all experimental variables except for the Configuration manipulation. Mean response times and standard deviations are summarized in table 2. All significant main effects were in the expected directions and no significant interactions were found. The extremely low error rates found in the data indicated that subjects followed a conservative strategy, i.e., they appeared to withhold judgment until they were confident of their interpretations, as they were encouraged to do by the experiment instructions. These low rates precluded a meaningful analysis of the error results and they were excluded from consideration. Results of the analysis of variance for mean response times are summarized in table 3. Complete data for cell means from the experiment are contained in the appendix.

Table 2. Response time results.

Manipulation	Level	Mean (sec)	SD
Number of Displays	1	5.96	0.9649
	2	9.89	1.5004
	3	11.61	1.9835
Number of Displayed Figures	1	6.97	0.9516
	2	10.59	1.7351
	3	14.01	2.5101
Number of Actual Figures	1	8.53	1.5448
	2	12.45	2.1944
	3	16.59	2.5224
Sensor Mobility	fixed	8.92	1.9312
	mobile	12.15	2.3625
Display Configuration	conv.	10.92	1.7981
	conform.	10.13	2.2408

Table 3. Analysis of variance: response time.

Source	SS	df	MS	F	p<
Number of Displays	100.553	2	50.277	126.312	.001
Error	3.981	10	3.981		
Number of Displayed Figures	148.936	2	74.468	95.685	.001
Error	7.783	10	0.778		
Number of Actual Figures	194.772	2	97.386	77.109	.001
Error	12.630	10	1.263		
Sensor Mobility	31.558	1	31.558	8.972	.001
Error	17.587	5	3.518		
Display Config.	1.848	1	1.848	0.752	n.s
Error	12.294	5	2.459		

EXPERIMENTAL VARIABLES

Number of Displays. A main effect for number of displays ($F(2,10) = 126.312$, $p < .001$) was obtained in the analysis of mean response times (means were 5.96, 9.89, and 11.61 seconds for one, two, and three displays, respectively). Providing the operator with a larger number of displays required a greater amount of time to interpret, as shown in figure 6. This rather "common sense" effect confirms the significant impact of the size of the display set on operator performance.

Number of Displayed Figures. A significant main effect for number of displayed figures ($F(2,10) = 95.685$, $p < .001$) showed that mean response times increased directly with the number of figures contained in the display set (figure 7). Mean response times for trials which presented one, two, or three figures on the console were 6.97, 10.59, and 14.01 seconds, respectively.

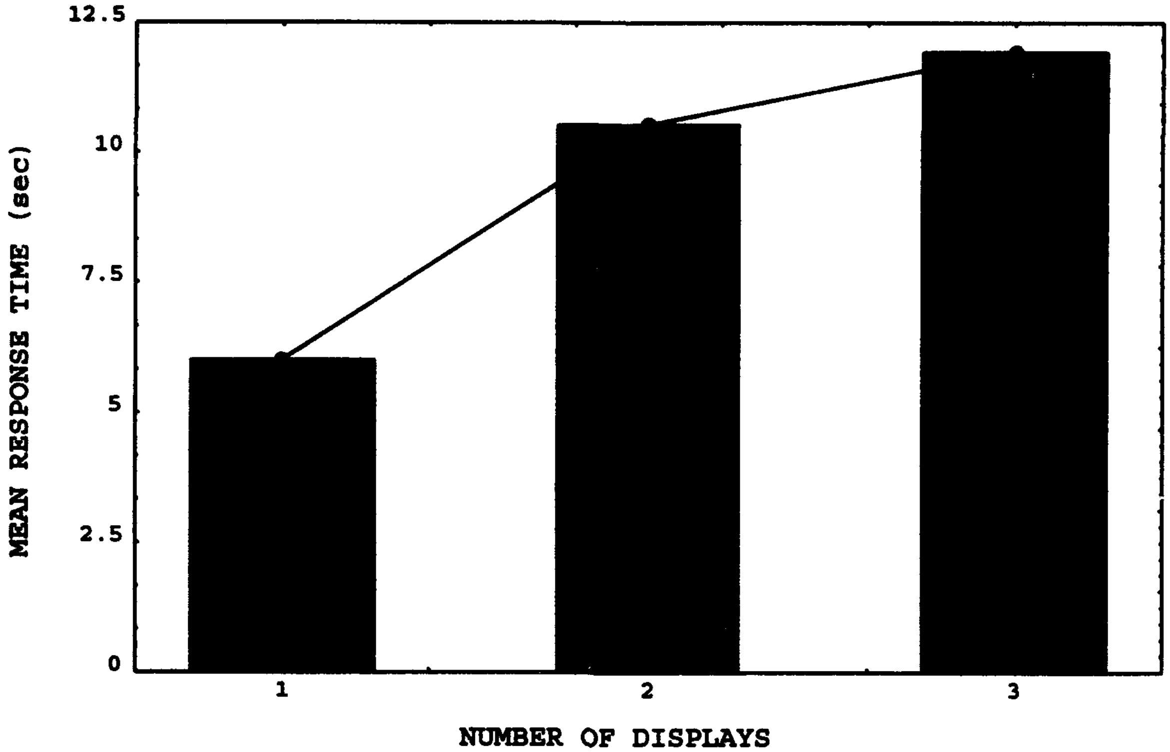

Figure 6. Mean response time versus number of displays.

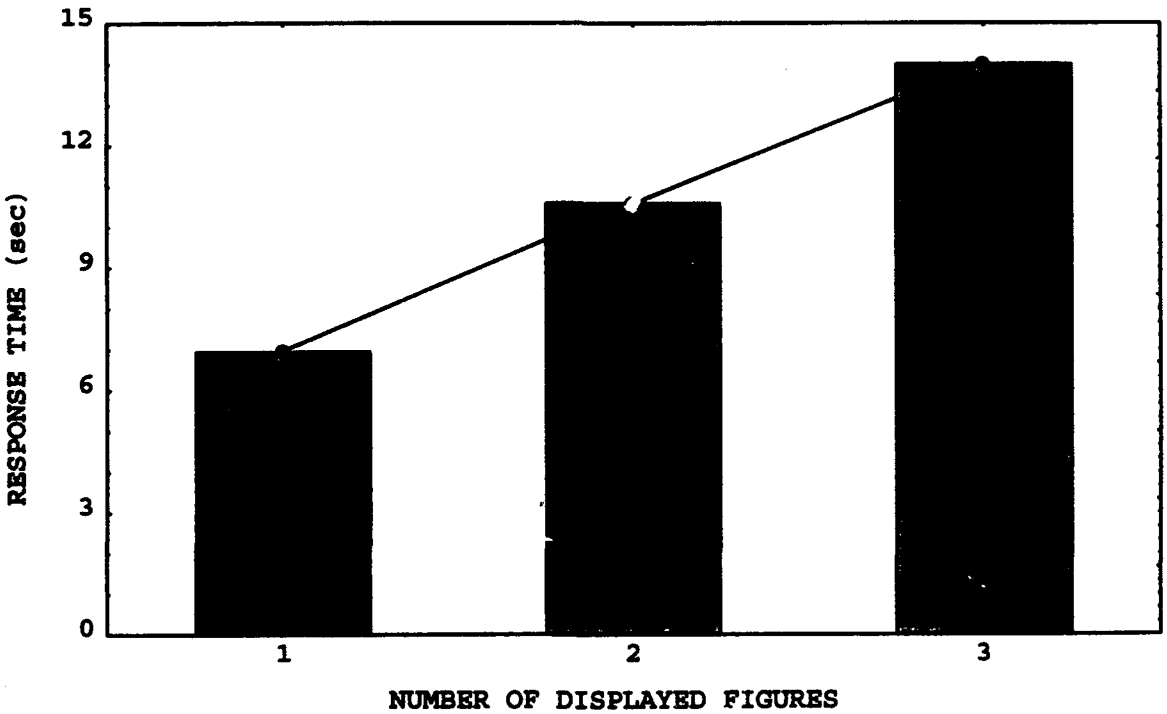

Figure 7. Mean response time versus number of displayed figures.

Number of Actual Figures. As shown in figure 8, a main effect for the number of actual figures was obtained for mean response times ($F(2,10) = 77.109$, $p < .001$). The conditions for this manipulation were labeled as follows: one involved total redundancy, i.e., only one figure was actually present in the building and was detected simultaneously by all sensors, generating a figure in each display. The three condition, by contrast, involved total independence, i.e., a unique figure was presented to each of three displays. The two condition was intermediate, depending on the display condition: if the trial involved two displays, then each figure was a unique intruder, but if the trial involved three displays, then two of them showed a common figure while the remaining display contained a unique intruder. The result appears to indicate that system operators take longer to process multiple independent images than they do to process redundant ones (mean response times were 8.53, 12.45, and 16.59 for one, two, and three actual figures, respectively).

Sensor Mobility. A significant main effect was obtained in the expected direction for the sensor mobility manipulation ($F(1,5) = 8.972$, $p < .001$). Interpretation of images from mobile sensor platforms took approximately 40 percent more time than interpretation of images from fixed-position platforms (mean response times were 8.92 and 12.15 seconds, for the fixed and mobile conditions, respectively). The results are shown in figure 9.

Display Configuration. The trend of mean response time data for the display configuration manipulation was in the expected direction (10.92 and 10.13 seconds, for the conventional and the conformal displays, respectively), but this effect was not statistically significant ($F(1,5) = 0.752$, $p < .434$).

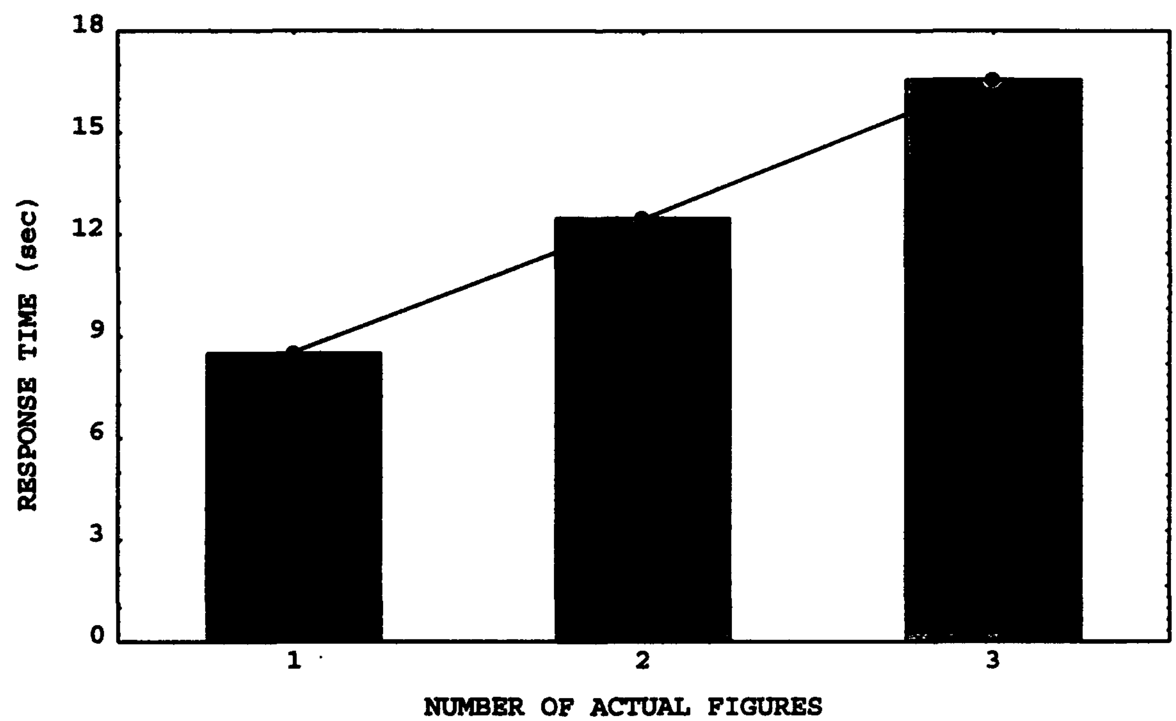

Figure 8. Mean response time versus number of actual figures.

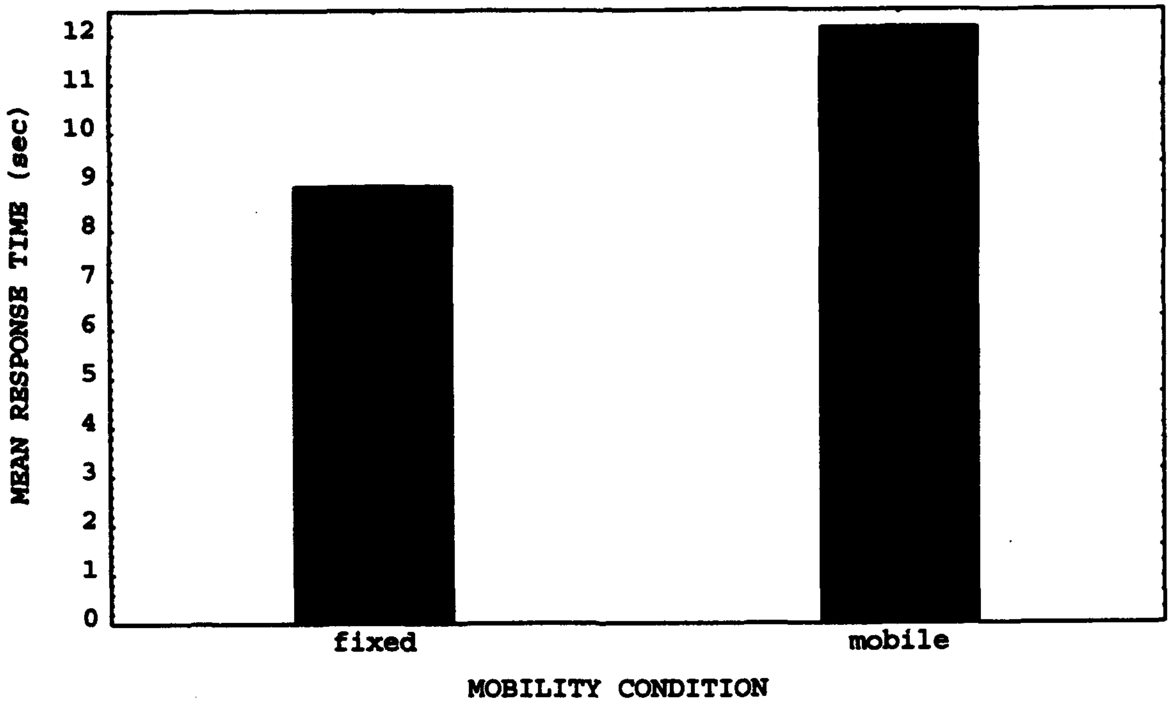

Figure 9. Mean response time versus sensor mobility.

DISCUSSION

In general, this experiment was successful in its goals of identifying and quantifying selected system design variables and their effects on human-machine performance. The sensitivity of operator performance to these manipulations and the confirmation provided for the major hypotheses were encouraging for such an exploratory effort. The study was also successful in identifying the nature and potential significance of variables needing further study.

The results generated by this experiment for size of display set and stimulus rate confirmed the original hypotheses about performance effects. Because these hypotheses were applications of existing perception and cognition theories, however, they add little new information to the human factors literature beyond a quantitative application of theory. Nevertheless, a major motivation for this work was to provide exactly such applications for system design guidance and in that sense, the results are worthwhile. Results from manipulating the number of actual figures in the display (i.e., an examination of image redundancy) are also sensible when the experience level of the subjects and the constrained nature of the images are considered, but the direction of this result was not predictable without the empirical test performed here. Further work can now be initiated with a clearer understanding of the effects of stimulus redundancy and operator experience on system performance.

The effect of platform mobility on response time is also sensible, but the magnitude of this result was not predictable prior to the study. Data from this manipulation is important because systems which include such autonomous, mobile sensors are currently undergoing advanced development and may find wide industrial application in the near future.

The results obtained from the use of a conformal display were not encouraging, and were most likely due to a combination of a small subject pool and a superficial application of the many perceptual and cognitive principles which bear on this problem. This topic deserves considerably more examination than was provided in this study.

RESPONSE TIME EQUATIONS

The data from this experiment were gathered to help develop design guidelines for systems using multiple remote platforms. Within the limits of the issues selected for study and the experiment design, it is possible to offer such guidelines in the form of regression equations which summarize the effects of the significant manipulations on response time. Such equations could provide a metric of expected performance with alternative design configurations, or could influence design approaches based on the expected system application. Thus, a system used in a low event rate environment might differ from one designed for high event rates.

Two equations were generated from the experiment data, using the stepwise regression technique of the STATISTICA software package. These equations reflect the response time performance from a fixed-platform system and a mobile-platform system because these two approaches showed such significant differences in the study.

Fixed-Platform Performance. The expected response time is formulated as:

RT (sec) = 2.440

+ 0.478 (Number of Actual Figures)

+ 0.328 (Number of Displayed Figures)

R^2 for this equation is 0.5157 (i.e., approximately 52% of the variance is accounted for by this expression). $F(2,117) = 62.762$, $p < .001$. The beta for Number of Displays was not significant at $p < .05$, so this term was not included in the equation.

Mobile-Platform Performance. The expected response time is formulated as:

RT (sec) = 2.035

+ 0.477 (Number of Displayed Figures)

+ 0.343 (Number of Actual Figures)

R for this equation is 0.5342. $F(2,117) = 67.083$, $p < .001$. Again, the beta for Number of Displays was not significant, and was not included in the equation. Further development of such expressions, to include additional design parameters or alternate measures of system effectiveness, would contribute to the human factors guidelines for multiple-platform control which motivated this study in the first place.

Detailed review of the methodological and theoretical characteristics of this study will be presented next.

METHODOLOGICAL ISSUES

The underlying experimental design proved to be troublesome, and represents the most serious issue for future improvement. As described in the INTRODUCTION, test conditions that showed multiple human figures in the same image were excluded from the design because their use might contaminate the experiment with undesired dimensions of task performance. This decision, however, resulted in an incomplete design (i.e., the shaded cells in table 1) and presented difficulties of statistical analysis, such as the inability to test for interactions. Further work in this topic must either refine the definitions of the underlying behaviors, to permit full factorial testing, or must use more sophisticated experimental designs to compensate for these difficulties.

The small subject pool was driven by conditions of the industrial environment in which this study took place. The engineering laboratory used for this research provided the system application, the equipment, and the environment for investigating an interesting and useful topic. It could not, however, provide access to subjects other than laboratory staff members; thus, a compromise with hard realities had to be made. The significance of the results indicated that this was not a severe limitation, but additional subjects are always nice to have and a larger pool might very well have modified the obtained results.

The data were fundamentally linear, or where they were not (e.g., response times for numbers of displays), there is evidence from the literature that they should be. In retrospect, the original decision about using small levels of these manipulations proved to be overly cautious. Important trends might have been detected if a greater range between manipulation levels had been used. If, for example, scan patterns or decision processes break down under significant task loads, then a larger range might have detected this effect. Further work in this area, therefore, should involve more careful selections of levels (e.g., one, five, and nine displays, etc.).

Response time proved to be an effective performance measure despite the noise inherent in the experimental equipment and protocol. Although manual entry of trial completion times was required by practical factors, subsequent research efforts will almost certainly include fully automated support for response time recording.

Measurement of operator errors (i.e., counting or location errors) was not effective, although this outcome was not unexpected. As described in the INTRODUCTION, system operators are most likely influenced by a criterion of accuracy, for reasons associated with doing an effective job. That is, the cost of a few seconds spent in resolving an alert situation can more than outweigh the costs of an erroneous response. In keeping with this perspective (possibly induced by the experiment instructions), subjects appeared to trade response time for some threshold level of accuracy. It should also be noted that the protocol for measuring location errors contained a large subjective component, as the experimenter had to judge response accuracy on the basis of pen marks made on many small paper diagrams. Subsequent apparatus for doing this work will most likely include some form of direct CRT-based position entry method, a reasonable requirement given that many current and emerging surveillance systems already have just such displays to show the layout of a patrolled area. It is possible that larger systems, involving more complexity or greater numbers of displays, may yet demonstrate significant error rates, in which case this measure will take on more practical importance. Future work in this area should therefore retain both time and accuracy measures.

THEORETICAL ISSUES

Number of Displays. The increase in response time as a function of the number of displays that had to be monitored was anticipated from the literature; the physical and perceptual tasks of orienting to an event require finite amounts of time.

According to the interpretation of the Hick-Hyman law described in the INTRODUCTION, the presence or absence of a figure in each image constitutes a binary decision involving one bit of information (i.e., $RT = \log_2 N$, where $N = 2$, representing the presence or absence conditions). Trials involving one display therefore require some unit of response time to process this one bit. Trials containing three displays, however, will require processing of 2.58 bits of data ($\log_2 N$, where $N = 6$, or three sets of binary choices). Assuming this model accurately captures the phenomenon, then the Hick–Hyman law would predict a linear relationship for this manipulation with a slope of 1.58 (i.e., 2.58 – 1.0).

In fact, a slope of 1.94 was obtained (i.e., the ratio of response time means of 11.61 and 5.96 seconds) which, within the accuracy limits of this study, is encouraging. Because the tasks involved in scanning are nominally the same for each display, however, the slope of response times should be linear — possibly more linear than the data showed (see figure 6). There is ample cause for caution, then, in interpreting these data according to information theory or any

other model. This is a case were additional data from a larger subject pool, from additional trials, or from additional levels of the manipulation would have been helpful.

Number of Displayed Figures. The effect of stimulus rate on response time was very linear, and lent support to predictions from the literature regarding increased information processing demands from increased stimulus loads.

There is a consistent increase in response time for each level of this manipulation. In addition, response times for one, two, and three displayed figures are uniformly higher than response times for one, two, and three displays. This indicates that a distinct amount of time is required to perform the image interpretation task, over the time required to orient to a display and perform the detection task. The mean response time elevation across the three conditions for Number of Displays and the Number of Displayed Figures manipulations is 1.37 seconds. The practical significance of these two manipulations is that both the number of displays, as well as their content, independently contribute to operator demands and increased performance times. In addition, the application of information theory to the task, in an effort to quantify the underlying processing demands, is very promising for further examination of the relative effects of both stimulus load and display design.

The lack of obtained asymptote to these manipulations, and the predicted linearity of their behavior from the literature, provide strong evidence that performance penalties would continue to accumulate, with obvious performance consequences for real world systems which are already planned with significantly higher numbers of platforms than those employed here.

Number of Actual Figures. This manipulation was originally conceived as a control for the complexity of displayed information. Scenes that presented more displayed images than there were actual figures in the environment were considered redundant, according to the definition used for this study. It was hypothesized from signal detection theory that redundant information could enhance the strength of a signal in noise by providing extra, mutually confirmatory cues to scene resolution if such cues were perceived by the operator as redundant (i.e., if their similarity were readily recognized). Because the primary dependent variable of this task was response time, one interpretation would be that greater redundancy would result in shorter response times, as the commonality between images was recognized. This was the obtained effect.

Unfortunately, practical constraints on the implementation of this study limited investigation to only one side of the issue. As discussed previously, this association between redundancy and signal strength would only be true for subjects who could make use of the redundant cues, and this was demonstrated with the use of trained and experienced subjects. This theory would also predict, however, that this redundancy might increase noise, rather than signal, for operators who were not familiar with the cues. The use of naive subjects, therefore, or use of multiple platform systems in unstructured environments, might produce results opposite to those found in this experiment. This is a primary issue for further research, and although this variable could not be completely addressed here, its outcome could have major implications for system design.

Although the trend in response times across conditions of redundancy fit the interpretation of signal detection theory provided earlier, the cell means were still quite high when compared to other manipulations. Clearly, the task of resolving redundancy generates additional processing demands.

Sensor Mobility. It was predicted that operator workload would be higher for monitoring tasks involving mobile systems than for systems which remained fixed in position. This hypothe-

sis was clearly borne out by the data; the flexibility and power of a semiautonomous system comes at a price in human performance. Furthermore, the extent of this problem was not fully resolved within the constraints of this study. It is quite possible that systems involving larger numbers of platforms, or unstructured settings, may generate even heavier penalties in response times or error rates.

It should be noted, however, that mobility was directly relevant to the operator task in this case. That is, the physical distribution of the platforms was a critical piece of information required to perform the task. It may be possible to redesign the task or to alter the nature of sensor cuing to mediate this problem, i.e., to create a situation where operator understanding of sensor distribution is not necessary to proper interpretation.

Display Configuration. The conformal display condition was included as an attempt to more fully and efficiently convey the information to operators. Specifically, sensor information and the geographic origin of that information were presented together in a single format. Because this approach provided more of the information required for task performance, it supposedly corresponded more closely with Wickens' (1984) discussions (chapter 2) of display-response matching, and it was predicted that performance would be improved under this condition. At the very least, it was believed that the conformal display would reduce the need to scan back and forth between a diagram of sensor positions and their video images, and that the conformal display would therefore reduce both the scanning time and short-term memory demands. Although the direction of the effect was in the predicted direction, the data did not reach statistical significance. It is possible that this design approach might be worth further study with a larger number of levels for other manipulations and with greater numbers of subjects. It is more likely, however, that the approach was too naive and that many more task parameters must be examined and addressed to generate meaningful display design support.

OTHER ISSUES

Sternberg (1969) subscribes to an additive model of multiple information processing stages and provided a statistical principle that could be used for analysis of studies such as the experiment reported here. Sternberg (1969) reasoned that two manipulations that influence a common processing stage will show an interactive effect on response time, but if separate stages are involved, there will only be additive effects. The complete absence of interactions for any of the manipulations for this experiment would support the position that the processing required to detect images in a display, and the processing required to interpret the meaning of those images, represent two distinct stages of the cognitive task. This extension of the information processing models discussed earlier is intriguing, but the limited data collected in this experiment, and the fact that this theory application is post hoc, would require that this kind of analysis be deferred to a future study.

A distinction was made earlier between systems which continually display sensor data and those which provide cued alerts to operators. Continual monitoring exacts a toll in human attention and vigilance, regardless of operator motivation, when events are of low frequency. An alerting system requires less operator monitoring, which would certainly make such a job more palatable. The offsetting price of this approach is that the operator must rapidly reorient himself or herself to the situation when an alert was given; it can be assumed that reliance on alerts would free operators to attend to other tasks between events, and therefore, attention is most likely directed elsewhere at the time of an alert.

Response time is a function of temporal uncertainty and expectancy (Wickens, 1984, chapter 9). If the observer knows when an event will occur, and can prepare for action, then response times are lower. These conditions held for the current experiment, i.e., observers were prepared for each trial. In real world applications, however, such conditions would not hold, indicating that even slower performance could be obtained. The results of this study apply to both types of system design: the number of displays is directly relevant to the continuous monitoring approach, while the number of displayed figures corresponds with the number of alerts in a semiautomated system. The distinction is important because changes to system parameters (e.g., number of sensors, expected frequency of environmental events, etc.) can have different performance effects on system output because two different types of cognitive tasks are involved.

FURTHER WORK

Of the four task characteristics for multiple platform control presented in the INTRODUCTION, only the second — determining the orientation and perspective of displayed sensor images, and developing an integrated spatial model — was examined in detail here. This was a reasonable beginning, as this requirement was essential to the other task characteristics in the list. Nevertheless, additional issues emerged from this experiment that should be resolved for near-term design guidance; other questions regarding system design still remain.

The most relevant research issues for near-term study involve the implications of operating in unstructured environments. The diminished predictability of scenes provided by video imagery from outdoor or complex environments, the possibility of greater background clutter, and the greater variety of potential target types further complicate the interpretive job of the operator. Although research literature is available on some of these issues (Cole & Pepper, 1986; Monk, 1976; Doll, et al., 1992), their applied effects in a multiple platform control setting, and their interactions with the manipulations studied here remain to be evaluated.

The images used for this experiment were static, i.e., little more than photographs of the sensed environment. Although it is likely that many systems will utilize such low bandwidth information, many current applications of multiple platform sensor systems (e.g., bank and factory security) impose no such restrictions, and real-time video imagery, with continuous motion, is part of the stimuli. Such motion can provide important cues and alerts to the activity being displayed. The extent of this enhancement, however, and the conditions under which it is best exploited, are not well documented.

The performance effects of alternative forms of information (principally from low-light level TV and infra-red sensors) also deserve study in the context of multiple platform control applications. Certainly, these sensors may represent sources of additional information to expedite image interpretation, much as the availability of redundant images may have reduced response time in the current study. Additional testing for this issue could greatly augment the tools available for system design, and protocols already exist for conducting such investigations efficiently (Kee, et al., 1992).

Finally, the ambivalent nature of the results for display configuration highlights the need to address this design issue in greater detail: what is the best method for indicating individual sensor perspective (whether a platform is pointing north, east, etc.) and for representing perspective relationships between platforms? Although not a planned part of this study, it was clear from observations and interviews of participants in this experiment that diagram information about

where each sensor was looking (i.e., in addition to the locations of sensors) would have been of significant help in integrating the image sets.

In summary, the dimensions of human perception and cognition represented by the task of multiple platform control are extremely rich, with applications to many forms of complex human-machine interaction. The results generated by this experiment should help to guide system interface design, but more importantly, should provide more specific direction for the additional research necessary to fully explore this topic.

REFERENCES

Ashby, W.R. 1956. *An Introduction to cybernetics.* London: Chapman & Hall, Ltd.

Fitts, P.M. 1966. "Cognitive aspects of information processing III: Set for speed versus accuracy." *Journal of Experimental Psychology*, vol. 71, pp. 849–857.

Gopher, D., and E. Donchin. 1986. "Workload—an examination of the concept." In *Handbook of Perception and Human Performance*, vol. II, pp. 41–1—41–49, K. Boff, L. Kaufman, & J. Thomas, Eds. New York: John Wiley.

Hick, W.E. 1952. "On the rate of gain of information." *Quarterly Journal of Experimental Psychology*, vol. 4, pp. 11–26.

Hyman, R. 1953. "Stimulus information as a determinant of reaction time." *Journal of Experimental Psychology*, vol. 45, pp. 423–432.

O'Brien, K. and E.M. Feldman. 1992. "Signal detection theory and methods for evaluating human performance in decision tasks." In *Sixth Annual Workshop on Space Operations, Applications, and Research*, pp. 489–496, K. Krishen, Ed., Springfield, VA: NASA/National Technical Information Service.

Pepper, R.L., and P.K. Kaomea. 1988. "Teleoperation: telepresence and performance assessment." Submitted to the Annual International Ergonomics Society Meeting, Teleoperation and Control, Bristol, England.

Scanlan, L.A. 1977. "Target acquisition in realistic terrain." In *Proceedings, 21st annual meeting of the Human Factors Society*, A.S. Neal and R. Palasek, Eds., Santa Monica: Human Factors Society.

Shannon, C.E., and W. Weaver. 1949. *The mathematical theory of communication.* Urbana, IL: University of Illinois Press.

Sheridan, T.B., and W.R. Ferrell. 1963. "Remote manipulative control with transmission delay." *IEEE Transaction on Human Factors in Electronics*, HFE–4, pp. 24–29.

Sternberg, S. 1969. "The Discovery of Processing Stages: Extensions of Donders' Method," *Acta Psychologica*, vol. 30, pp. 276–315.

Vickers, D. 1970. "Evidence for an accumulator model of psychophysical discrimination." *Ergonomics*, vol. 13, pp. 37–58.

Wickens, C. 1984. *Engineering Psychology and Human Performance*. Columbus, OH: Charles E. Merrill.

Wickens, C.D., S.J. Mountfor, and W.S. Schreiner. 1982. "Time-sharing efficiency: evidence for multiple resources, task hemispheric integrity and against a general ability." *Human Factors*, vol. 23, pp. 211–229.

Wickens, C.D., D.L. Sandry, and M. Vidulich. 1983. "Compatibility and resource competition between modality of input, central processing, and output: testing a model of complex task performance." *Human Factors*, vol. 25, pp. 227–248.

Woods, D.D. 1988. "Coping with complexity: The psychology of human behavior in complex systems." In *Mental Models, Tasks and Errors*, L.P. Goodstein, H.B. Andersen, and S.E. Olsen, Eds. London: Taylor & Francis.

APPENDIX

INSTRUCTIONS TO PARTICIPANTS

"You should imagine yourself as the operator of an industrial security monitoring station. Your job is to manage a bank of displays that show scenes from around a factory, bank, laboratory, etc. and to call for a security response if one is needed.

You are going to see a group of displays, which represent the kind of video camera information you would see from a guard station, located in the plant. You may have to monitor one, two, or three of these displays at the same time, representing the views of one or more cameras. Displays will come on in front of you, which represent the kind of information you would see if a remote security station had set off an alert. These alerts may or may not all be valid; that is, they may have been triggered by an erroneous signal, so you will have to examine all the displays to determine what is going on.

You are looking for an intruder. Although some of the displays may not contain an image of an intruder, there will always be at least one intruder somewhere on the monitor; if one display is empty, another display will have something in it. There may be more than one intruder. There may also be more than one picture containing an image of the same intruder. That is, camera views can overlap, giving you a picture of the same person from different angles. Your job is to figure out how many intruders are in the plant, and where they are located.

When you tell me you are ready, I will hand you a diagram that shows you where the sensor cameras are located in the building. I will also put an image, or a set of images, on the screen which have been transmitted from these sensors. Based on these images, you must look at the monitor and place "Ks" on the diagram to show where you think the intruders are. When you are done, lift your pencil off of the paper and lean back; when I see you do this, I will stop the trial. Your performance is being timed.

You may have fixed or mobile sensor platforms. That is, each trial can come from sensors that are fixed in their locations around the building or it can come from sensors which have changed their positions from the previous trial. I will always tell you which condition you are operating in, and the diagram I give you will show you where the platforms are. For this session, you will be dealing with (fixed/mobile) sensor.

You may also see one of two kinds of displays, either a conventional set of CRT's set in a bank and numbered left to right, or a set which is located on the monitor in the same way they are distributed in the building. In the moveable condition, you should notice that each video image is located on the screen in the same relative location that you see it on the diagram. I will always tell you which kind of system you will have. For this session, you are going to see (conventional/conformal) displays only.

You may have one, two, or three displays to monitor, as if you were responsible for one, two or three sensor cameras. I will always tell you what you can expect. For this session, you will see (one/two/three) images only.

Do you have any questions? Again, for this set of trials, you will be looking at (repeat system conditions)."

Cell means for repeated measures tests (sec)

SUBJ.	NUMBER DISPLAYS			NUMBER OF DISPLAYED FIGURES			NUMBER OF ACTUAL FIGURES			SENSOR MOBILITY		DISPLAY CONFIGURATION	
	1	2	3	1	2	3	1	2	3	fixed	mobile	conv.	conform.
1	5.90	9.53	12.55	7.26	10.93	14.77	8.78	12.84	18.60	10.55	11.41	10.86	11.10
2	4.95	8.20	9.66	5.61	8.77	11.88	6.62	10.73	15.60	8.60	8.91	8.00	9.51
3	5.48	9.28	10.76	7.19	9.79	12.37	8.69	10.26	14.05	6.12	13.46	9.95	9.62
4	5.55	9.41	9.73	6.35	9.40	11.83	7.45	11.31	13.48	7.28	11.15	11.33	7.10
5	6.13	10.29	12.09	6.94	10.97	14.95	8.45	13.39	18.68	9.85	12.06	12.31	9.60
6	7.75	12.63	14.85	8.44	13.66	18.26	11.18	16.19	19.10	11.04	15.91	13.07	13.88

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*

2. REPORT DATE
July 1994

3. REPORT TYPE AND DATES COVERED
Final

4. TITLE AND SUBTITLE
HUMAN PERFORMANCE STUDIES FOR CONTROL OF MULTIPLE REMOTE SYSTEMS

5. FUNDING NUMBERS
PE: 0602234N
WU: DN303208
Subproj: R3420

6. AUTHOR(S)
Steven A. Murray

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Naval Command, Control and Ocean Surveillance Center (NCCOSC)
RDT&E Division
San Diego, CA 92152–5001

8. PERFORMING ORGANIZATION REPORT NUMBER
TR 1658

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution is unlimited.

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Many systems involving multiple remote sensors or machines require a single operator to control more than one device simultaneously. Interface design issues for operator support in such systems can be formidable, as the inherent task complexity creates significant opportunities for operator confusion and overload. A study was conducted to examine some of these issues, using a simulated industrial security environment as the applied setting. The operator task was to designate the number and location of intruders in a simulated building, using video information from remote sensor platforms. The experiment manipulated the number of displays which had to be monitored, event rate, image redundancy and sensor platform mobility. Response time increased significantly for increasing numbers of displays, as expected, but also increased independently for event rate image redundancy and for mobile sensors. Results showed that significant performance penalties may be encountered in multiple platform control, and that these penalties accumulate at seemingly low levels of complexity.

14. SUBJECT TERMS
Remote systems
Human performance

15. NUMBER OF PAGES
44

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	SAME AS REPORT

UNCLASSIFIED

21a. NAME OF RESPONSIBLE INDIVIDUAL	21b. TELEPHONE (Include Area Code)	21c. OFFICE SYMBOL
Steven A. Murray	(619) 553-6350	Code 531

NSN 7540-01-280-5500

Standard form 298 (BACK)

UNCLASSIFIED

INITIAL DISTRIBUTION

Code 0012	Patent Counsel	(1)
Code 0274	Library	(2)
Code 0275	Archive/Stock	(6)
Code 50	H. O. Porter	(1)
Code 53	D. W. Murphy	(1)
Code 531	D. E. Demuth	(1)
Code 531	Steve A. Murray	(25)

Defense Technical Information Center
Alexandria, VA 22304–6145 (4)

NCCOSC Washington Liaison Office
Washington, DC 20363–5100

Center for Naval Analyses
Alexandria, VA 22302–0268

Navy Acquisition, Research and Development
Information Center (NARDIC)
Arlington, VA 22244–5114

GIDEP Operations Center
Corona, CA 91718–8000